iBeCOOL – Extraction d'informations biographiques dans les textes financiers

Michaela Geierhos[1], Olivier Blanc[2], Sandra Bsiri[1]
Université Ludwig-Maximilian Munich, Université Catholique de Louvain

Résumé

Ce papier présente le contexte linguistique et la modélisation de notre système iBeCOOL (Informations Biographiques Extraites à l'aide de COntextes Observés Linguistiquement) dédié à l'extraction d'informations biographiques dans les textes de la presse financière en langue anglaise. La notion d'événement biographique (tel que la naissance, le mariage, la carrière professionnelle) est caractérisée formellement par un schéma prédicatif à plusieurs arguments dont l'un étant une instance de la classe d'objets ⟨Personne⟩. Notre approche consiste à décrire ces types de relations à l'aide de grammaires locales et de lexiques terminologiques. Nos résultats montrent que cette approche semble viable et nous poussent à élargir cette étude par l'analyse de nouveaux genres textuels.

Mots-clés : extraction d'informations biographiques, relations sémantiques, grammaires locales, entités nommées, enrichissement du lexique.

1. Introduction

La masse toujours croissante de documents disponibles sur Internet rend l'accès à l'information difficile même avec l'usage indispensable des moteurs de recherche. Les informations relatives aux personnalités qui font l'actualité dans le monde ne font pas exception. Il est, en effet, très rare de trouver des biographies complètes sur les personnes d'intérêt public. Les textes journalistiques donnent de nouvelles informations biographiques concernant les personnes citées, mais celles-ci demeurent néanmoins souvent partielles. Ainsi pour répondre à un besoin informationnel sur un personnage de notoriété mondiale, l'utilisateur est contraint de lancer une requête sur un moteur de recherche à partir du nom de la personne et de mots clés associés à l'événement biographique d'intérêt, de parcourir ensuite les multiples documents résultats, susceptibles de contenir l'information recherchée, pour enfin rassembler les différents passages associés à son besoin et retracer ainsi la biographie souhaitée. Cette méthode s'avère répétitive et peu fiable, car l'information présente dans les diverses sources électroniques n'est pas indexée sur des critères sémantiques par les moteurs de recherche mais les documents sont prétraités uniquement sur la base des mots qu'ils contiennent. C'est donc à l'utilisateur que revient la tâche de classer les différents documents donnés en réponse par le système et de retrouver dans ceux-ci les segments pertinents pour ses besoins personnels.

[1]Centrum für Informations- und Sprachverarbeitung, Ludwig-Maximilians-Universität München – {michaela.geierhos,sandra.bsiri}@cis.uni-muenchen.de
[2]CENTAL, Université Catholique de Louvain – olivier.blanc@uclouvain.be

L'extraction des données biographiques contenues dans les diverses ressources textuelles non structurées est une tâche complexe, fortement dépendante de la langue et des phénomènes linguistiques associés (Poibeau 2003). De nombreux auteurs ont reconnu la complexité du problème (Kanzaki 2007; Davis & Galbraith 2004) et se sont intéressés à l'identification des indicateurs spécifiques internes et externes utiles à la reconnaissance des événements biographiques (Agichtein & Gravano 2000; Duboué *et al.* 2003).

Les travaux présentés ici se concentrent essentiellement sur le sous-langage économique et financier de la presse financière en langue anglaise. La performance des extractions est d'autant plus satisfaisante que les documents analysés se restreignent à ce domaine spécialisé. Cette tâche peut être utile pour plusieurs domaines applicatifs distincts comme les systèmes de résumés biographiques à partir de plusieurs documents (Schiffman *et al.* 2001), ou encore les systèmes de question-réponse (Tsur *et al.* 2004).

La notion d'information biographique est définie dans la seconde section de cet article, dans laquelle nous détaillons les entités susceptibles d'intervenir dans de telles relations. Dans les sections 3 et 4 nous présentons les ressources linguistiques (grammaires locales et dictionnaires terminologiques) que nous avons construites pour l'extraction de ce type de relations. Nous présentons nos méthodes utilisant des patrons lexico-syntaxiques, pour l'acquisition semi-automatique de nouvelles instances terminologiques et la découverte de locutions verbales synonymiques à partir de verbes initiaux qui nous ont permis d'enrichir de manière conséquente ces deux types de ressources. Nous concluons en section 5 et 6 en donnant une évaluation de la qualité d'extraction de notre système.

2. L'information biographique

Retracer la biographie d'une personne consiste à établir la liste de l'ensemble des événements qui sont survenus tout au long de sa vie, tels que sa naissance, son parcours scolaire et professionnel, ses relations privées et professionnelles établies avec d'autres individus, sa mort. Pour délimiter formellement quels sont les énoncés en langue naturelle dont la sémantique porte une information biographique, nous utilisons le modèle des classes d'objets (Gross 1994; Le Pesant & Mathieu-Colas 1998) : Les classes d'objets sont « des classes sémantiques construites à partir de critères syntaxiques » ; celles-ci sont définies par des prédicats définitionnels sémantiquement homogènes de type verbes, adjectifs ou noms auxquels correspondent des domaines d'arguments. Par exemple, la classe ⟨Profession⟩ étudiée par (Buvet & Foucou 2001), est l'ensemble des noms simples et composés répondant essentiellement aux prédicats *« gagner sa vie comme »* et *« exercer la profession de »* et contient des instances telles que *« ingénieur, instituteur, second de cuisine »*.

Dans ce contexte, nous définissons une information biographique comme une relation prédicative à plusieurs arguments dont l'un est une entité de la classe ⟨Personne⟩. Il n'y a pas de restriction de sélection sur les autres intervenants de la relation pour que celle-ci portent une information biographique. Cependant, dans les différentes relations que nous avons étudiées dans le cadre de ces travaux, les autres arguments sont typiquement des instances de classes ⟨Personne⟩, ⟨Lieu⟩, ⟨Date⟩, ⟨Organisation⟩, ⟨Branche⟩, ⟨Matière⟩ ou ⟨Profession⟩.

L'« information biographique » est ainsi une relation entre une personne et un événement biographique, qui peut être exprimée à l'aide d'un verbe prédicatif, retraçant le portrait de quelqu'un (1a–1c).

(1) a. *Sigman, born in Brooklyn in 1909.*
   b. *Andrew Gilligan graduated from Cambridge with a degree in history.*
   c. *Jim Sweeney will also be joining AmeriQuest as Vice President.*

La sémantique d'événement biographique n'est pas nécessairement portée par un verbe mais peut être portée par d'autres catégories grammaticales prédicatives comme le montrent les phrases à verbes supports suivantes :

(2) a. *Elizabeth gave birth to a little girl in May 2004.*
b. *Paul and Claire became man and wife in 1998.*
c. *John Smith was born in Florida on February 12, 1965.*
d. *Jacob McCandles is six-feet under.*
e. *Gov. Greenhalge breathed his last at his home in Lowell, Mass.*

Ces exemples montrent que la fonction de prédicat peut aussi être remplie par un nom prédicatif (2a–2b), un adjectif prédicatif (2c) ou une expression idiomatique (2d–2e), nous parlons dorénavant de locution verbale pour désigner n'importe lequel de ces types de prédicats.

Dans le cadre de ces travaux, nous nous sommes intéressés à 12 types d'événements biographiques : 6 événements que nous classifions dans les événements personnels : la naissance, l'enfance, la formation, le mariage, le divorce et le décès ; 6 événements relatifs à la carrière professionnelle d'une personne : l'obtention d'un emploi, l'occupation d'un poste, le licenciement, le succession, la démission et le départ en retraite.

3. Les grammaires locales et sous-langages

Nous partons de l'hypothèse que l'ensemble des énoncés décrivant une information biographique se caractérisent par un lexique de taille finie et un nombre de schémas de phrases limité ; cette hypothèse est d'autant plus vérifiée que nous nous limitons à traiter des textes d'un genre particulier (ici, la presse anglo-saxonne économique et financière). En ce sens, nous considérons qu'ils constituent un sous-langage, dans le sens de (Harris 1968). (Hunston & Sinclair 2000) montre qu'il est possible de considérer les grammaires locales comme des petits sous-langages et que par conséquent, pour un domaine donné (ici, l'information biographique), il est possible d'élaborer un ensemble de grammaires locales étendu couvrant au mieux la totalité du sous-langage.

Les grammaires locales que nous considérons sont des réseaux de transitions récursifs, représentés par des graphes dont la construction et la manipulation sont facilitées par le logiciel libre `Unitex`[3] (Paumier 2004; Silberztein 1993). Elles n'ont pas pour vocation de décrire l'ensemble de la grammaire d'une langue, mais décrivent les structures syntaxiques et lexicales des phénomènes linguistiques propres à ce langage spécialisé (Nakamura 2005).

4. Enrichissement du lexique à travers les contextes riches en connaissances

Nous définissons précédemment la notion d'« information biographique » comme une relation prédicative entre plusieurs arguments dont l'un est une entité de la classe ⟨Personne⟩ et les autres intervenants pouvant alternativement appartenir aux classes d'objets ⟨Lieu⟩, ⟨Date⟩, ⟨Organisation⟩, ⟨Branche⟩, ⟨Matière⟩ ou ⟨Profession⟩. Chacune de ces classes a été traduite par un dictionnaire électronique de la forme DELA (Courtois & Silberztein 1990; Courtois 2004) dont les entrées lexicales reprennent les instances hyponymes du nom de la classe d'objets source. La pertinence de l'analyse contextuelle et par conséquent la qualité du système d'extraction automatique des faits biographiques est d'autant plus satisfaisante que l'on dispose de bases de connaissances riches en entrées lexicales. Ainsi, plus les classes

[3] `http://www-igm.univ-mlv.fr/~unitex`

d'objets sont riches en instances, plus les extractions sont pertinentes et les analyses contextuelles nécessaires à la levée d'ambigüité sont rudimentaires.

4.1. Acquisition automatique de nouvelles entrées lexicales

Nous avons ainsi constitué un lexique de spécialité propre au monde de l'entreprise (activité professionnelle, secteur d'activité, etc.), ainsi qu'un lexique de noms propres pour les personnes, les toponymes et les organisations à partir de diverses ressources disponibles sur Internet (Wikipédia, WordNet, Biography.com, SpecialistInfo.com, ZoomInfo.com, Guide to the World of Occupations[4], LabourMarket[5], MapPlanet.com, Occupational Outlook Handbook[6], Prospects.ac.uk, etc.) et des lexiques du laboratoire CIS. Cependant, lors de nos tests préliminaires d'analyse sur corpus, nous avons observé qu'il existait de nombreuses unités lexicales, hyponymes des classes d'objets présentées ci-dessus qui n'étaient pas encore recensées dans nos dictionnaires électroniques. C'est là qu'interviennent les patrons lexico-syntaxiques ou « knowledge patterns » (KPs) (Meyer 2001) pour découvrir de nouvelles instances associées à ces différentes catégories. Nous montrons sur un exemple de la classe ⟨Secteur d'activité⟩ comment un tel gain automatique de connaissance est possible. La désignation du secteur d'activité respecte une structure syntaxique où le contexte droit est souvent représenté par un descripteur comme *« industry »*, *« sector »* ou *« company »* qui peut agir en qualité de déclencheur pour reconnaître de nouveaux noms de secteurs :

(3) a. *administration sector*
b. *automobile industry*
c. *arts and leisure sector*

Ainsi un KP recensant ces descripteurs et décrivant la structure syntaxique interne d'un nom de secteur permet d'identifier de nouvelles instances de cette classe dans les textes analysés. Les noms des secteurs de l'exemple (3) peuvent être également retrouvés dans les textes accompagnés par des descripteurs différents : il ne s'agit pas de séquences figées comme le montre le terme *automobile* de l'exemple (3b) qui peut aussi apparaître dans d'autres séquences comme *« automobile industry»* ou *« automobile business »*. Pour ce genre de termes, nous ne retenons dans les classes d'objets correspondantes que la séquence sans le descripteur. Ce qui revient ici à ajouter le terme *« automobile »* aux instances de la classe ⟨Secteur d'activité⟩ et non pas les deux séquences *« automobile industry »* et *« automobile business »* qui seront reconnues ultérieurement dans les textes à travers les grammaires descriptives dotées des déclencheurs *« business »* et *« industry »*. Certains noms de secteurs sont cependant moins flexibles ; c'est notamment généralement le cas des termes associés au descripteur *« service »*, tels que *« reparation service »* ou *« animal physiotherapy services »*. De telles séquences, où *« service »* ne peut être remplacé par aucun autre descripteur, viennent enrichir les classes d'objets en tant que mots composés figés.

Cette méthode nous a permis de découvrir plus de 40 000 nouvelles entrées à partir de 10 000 instances initiales pour la classe ⟨Secteur d'activité⟩ vérifiées manuellement. Le tableau présenté en annexes résume l'ensemble des classes d'objets que nous avons identifiées comme étant a priori nécessaires au bon déroulement de l'extraction de l'information biographique. La somme des entrées dépasse, à ce jour, les 10 millions d'unités lexicales dont environ 15 % étaient découvertes à l'aide de nos méthodes de Bootstrapping. Chaque unité est représentée dans le dictionnaire, accompagnée d'une étiquette sémantique correspondante au nom de la classe d'objets associée.

[4]http://www.occupationsguide.cz/en/abecedni/abecedni.htm
[5]http://www.labourmarket.co.nz/labourmarket.htm
[6]http://www.umsl.edu/services/govdocs/ooh20002001/1.htm

4.2. Bootstrapping et extraction des prédicats synonymiques

Pour chacun des 12 types d'événements biographiques étudiés nous avons sélectionné une liste restreinte de verbes en relation à partir desquels nous avons extrait des locutions verbales synonymiques de manière semi-automatique.

La découverte des relations synonymiques associées aux verbes initiaux fut effective à l'aide de techniques de Bootstrapping (Gross 1999) appliquées de manière itérative. Le processus complet d'extraction de candidats synonymiques est assez complexe et implique un enrichissement des contextes internes par la détection de contextes externes spécifiques et inversement. Pour une description complète de l'ensemble du processus nous renvoyons à (Geierhos 2007).

Nous avons ainsi recensé un total de 145 locutions verbales ayant une sémantique en rapport avec l'une des 16 relations biographiques étudiées (cf. section 2) ; parmi celle-ci, 50 prédicats sont relatifs à des informations dites personnelles (telles que la naissance, le mariage, etc.), et 95 prédicats concernent les informations professionnelles, mettant en relation un individu avec une entreprise, une date d'embauche, de départ, une activité exercée, etc.

5. Évaluation des schémas prédicatifs

Afin d'évaluer la qualité d'extraction des schémas prédicatifs décrits dans notre grammaire, nous avons annoté manuellement un corpus de test de 4 500 phrases constitué d'articles du quotidien Financial Times du mois de juin 2004. Nous avons choisi ces articles de telle sorte que le maximum des prédicats étudiés soient représentés.

Le tableau 1 résume les résultats obtenus pour l'extraction des différentes informations professionnelles. On peut y lire une précision moyenne atteinte de 96.3 % et un rappel moyen de 87.7 %.

Schéma initial	**Précision**	**Rappel**
X was appointed as P	96.9 %	92.1 %
X joint O as P (of B) (in D)	97.8 %	88.4 %
X was employed as P	94.1 %	87.9 %
X was payed as P (by O)	91.6 %	88.0 %
X worked as P (for O)	96.2 %	90.6%
X was dismissed as P	95.8 %	89.5 %
X was replaced as P (by Y)	98.4 %	95.6 %
X resigned as P of O	98.8 %	94.0 %
X retired as P (in D)	97.2 %	63.3 %
En moyenne	96.3 %	87.7 %

TAB. *1. Résultats d'évaluation sur le corpus de test*

On peut par exemple observer que les variations syntaxiques du prédicat *« X retired as P (in D) »* sont loin d'être complètement décrites puisque nous obtenons un rappel de 63.3 % pour ces constructions. En effet la grammaire correspondante à ce prédicat reconnaît des phrases comme :

```
<FN>Ian</FN> <SN>McLeish</SN> retired as <JD>manager</JD> with effect from
<DATE>31 December 2003</DATE>.
```

mais aussi des segments de phrases, où seule une partie de l'information recherchée est reconnue comme dans la concordance :

```
David Selwood, 69, who retired as <JD>resident judge</JD> of
<GEO>Portsmouth</GEO> crown court last week on health grounds.
```

Ces concordances partiellement reconnues sont prises en compte pour le calcul du rappel et de la précision comme étant des résultats négatifs. Du fait de l'insertion de l'âge entre l'entité ⟨Personne⟩ et la phrase relative contenant le prédicat étudié, la reconnaissance n'a pas été possible car cette forme syntaxique n'avait pas été apprise antérieurement. A cet état, on peut justifier notre choix de l'utilisation des grammaires locales pour la modélisation de nos patrons d'extraction, méthode qui nous permet d'avoir des règles visiblement lisibles et compréhensibles et qui facilite par conséquent la maintenance et la mise à jour de nos grammaires.

Une possibilité d'améliorer ces résultats est de construire des grammaires indépendantes des prédicats synonymes des constructions verbales initiales et dont la structure syntaxique décrite actuellement se limite à celle des prédicats initiaux. A ce jour, seule la variante syntaxique de type *« X served as P »* est modélisée comme sous-grammaire du prédicat initial *« X worked as »*. Pour aspirer à une couverture plus large des structures syntaxiques et sémantiques du verbe *« to serve »*, il est indispensable de lui consacrer une grammaire propre. C'est pourquoi il n'est pas étonnant que l'évaluation d'un tel prédicat sur le corpus de test atteigne une précision de 97.1 %, semblable à celle du prédicat *« to work as »*, mais que le rappel en revanche soit très bas et n'atteigne que les 33.8 %. Cette observation est valable pour tous les prédicats synonymiques acquis par Bootstrapping dans la phase d'apprentissage. Ce résultat n'est en aucun cas une surprise, il confirme seulement que les prédicats sémantiquement semblables usent de structures syntaxiquement différentes pour paraphraser un même événement.

Dans une perspective d'amélioration des performances de notre système, nous nous sommes ainsi lancés dans le développement d'autres grammaires autonomes pour les structures prédicat-argument synonymiques.

Notons que ces chiffres ne donnent pas une évaluation exacte de la qualité d'extraction de notre système sur les textes de la presse financière en général, puisque le corpus de test a été préalablement constitué en fonction de la présence de prédicats prédéfinis dans notre système. En ce sens nous évaluons ici la qualité d'extraction de ces prédicats déjà définis et non la qualité d'extraction de l'information biographique en général. En l'absence de corpus de référence annotés manuellement par des relations biographiques, il nous est très difficile de faire une telle évaluation, qui serait par ailleurs plus pertinente du point de vue applicatif.

6. Conclusion et perspectives

L'extraction automatique de l'information biographique est une tâche complexe qui dépend fortement des connaissances acquises sur le domaine à analyser. Nous avons montré dans cet article l'intérêt des grammaires locales comme formalisme de représentation des variabilités syntaxiques existantes entre les instances des classes d'objets et nous avons présenté notre processus itératif basé sur l'utilisation de patrons lexico-syntaxique et des méthodes de Bootstrapping utiles pour la reconnaissance et l'extraction d'information dans les textes ainsi que pour l'acquisition automatique de terminologie nouvelle.

Nos résultats préliminaires nous incitent à poursuivre notre étude en l'étendant à la description d'autres événements biographiques et surtout à l'analyse de nouveaux genres textuels afin de découvrir de nouvelles locutions verbales synonymiques nous permettant de réaliser à long terme un système capable de

générer automatiquement la biographie d'une personne en extrayant l'information pertinente dispersée dans une multitude de documents.

Références

AGICHTEIN E. et GRAVANO L. (2000), "Snowball : Extracting Relations from Large Plain-Text Collections", in *Proceedings of the Fifth ACM International Conference on Digital Libraries*,San Antonio, Texas, USA : 85–94.

BUVET P.-A. et FOUCOU P.-Y. (2001), "Classes d'objets et recherche sur le web", in *Lingvisticæ Investigationes*, vol. 23.

COURTOIS B. (2004), "Dictionnaires électroniques DELAF anglais et français", in et Éric Laporte et Mireille Piot et Max Silberztein C. L. (Ed), *Lexique, syntaxe et lexique-grammaire ; syntax, lexis & lexicon-grammar*, John Benjamins, Amsterdam/Philadelphia : 113–123.

COURTOIS B. et SILBERZTEIN M. (1990), "Dictionnaires électroniques du français", in *Langues française*, vol. 87.

DAVIS I. et GALBRAITH D. (2004), *BIO : A vocubulary for biographical information*, http://purl.org/vocab/bio/.

DUBOUÉ P., MCKEOWN K. et HATZIVASSILOGLOU V. (2003), "ProGenIE : Biographical descriptions for Intelligence Analysis", in *Proceedings of the NSF/NIJ Symposium on Intelligence and Security Informatics*, vol. 2665,(Lecture Notes in Computer Science) : Springer, Tucson, Arizona, USA : 343–345.

GEIERHOS M. (2007), *Grammatik der Menschenbezeichner in biographischen Kontexten*, Rapport Technique, Centrum für Informations- und Sprachverarbeitung (CIS), Ludwig-Maximilians-Universität, Munich, Allemagne.

GROSS G. (1994), "Classes d'objets et description des verbes", in *Langages*, vol. 115.

GROSS M. (1999), "A bootstrap method for constructing local grammars", in *Contemporary Mathematics : Proceedings of the Symposium, University of Belgrad*,Belgrad, Serbie : 229–250.

HARRIS Z. S. (1968), *Mathematical Structures of Language*, John Wiley & Sons, New York.

HUNSTON S. et SINCLAIR J. (2000), "A local grammar of evaluation", in Hunston S. & Thompson G. (Eds), *Evaluation in Text : authorial stance and the construction of discourse*, Oxford University Press, Oxford, England : 74–101.

KANZAKI M. (2007), *Who's who description vocabulary*, http://www.kanzaki.com/ns/whois.

LE PESANT D. et MATHIEU-COLAS M. (1998), "Introduction aux classes d'objets", in *Langages*, vol. 131.

MEYER I. (2001), "Extracting knowledge-rich contexts for terminography", in Bourigault D., Jacquemin C. & L'Homme M.-C. (Eds), *Recent Advances in Computational Terminology*, John Benjamins, Amsterdam : 279–302.

NAKAMURA T. (2005), "Analysing Texts in a Specific Domain with Local Grammars : The Case of Stock Exchange Market Reports", in *Linguistic Informatics – State of the Art and the Future*, vol. 1.

PAUMIER S. (2004), *Manuel d'utilisation d'Unitex*, http://wwwigm.univmlv.fr/~unitex/.

POIBEAU T. (2003), *Extraction automatique d'information, du texte brut au web sémantique*, Lavoisier.

SCHIFFMAN B., MANI I. et CONCEPCION K. J. (2001), "Producing Biographical Summaries : Combining Linguistic Knowledge with Corpus Statistics", in *Proceedings of the* 39th *Annual Meeting of the Association for Computational Linguistics*,Toulouse, France : 450–457.

SILBERZTEIN M. (1993), *Dictionnaire électroniques et analyse automatique de textes - Le systèm e INTEX*, Paris, Masson.

TSUR O., DE RIJKE M. et SIMA'AN K. (2004), "BioGrapher : Biography Questions as a Restricted Domain Question Answering Task", in *Proceedings ACL 2004 Workshop on Question Answering in Restricted Domains*,Barcelone, Espagne : 23–30.

Annexe : Récapitulatif des différentes classes d'objets recensées

Classe d'objet	***Sous-classe d'objet***	***Balise sémantique***	***Nombre d'instances***	***Exemples d'instances***
Nome Propre	Titre	⟨Title⟩	370	*Queen, Lord, PhD, Mr.*
	Prénom	⟨FirstName⟩	38 500	*Lara, Marie-Luise, Ben*
	Nom de Famille	⟨Surname⟩	1 250 000	*Oltay-Smith, Yildiz*
	Nom de Personne (Complet)	⟨LongName⟩	8 300 000	*Henna Nordqvist*
Rôle Social	Famille	⟨Human⟩	6 400	*daughter, son, aunt*
	Profession	⟨JobDescriptor⟩	45 000	*cook, kitchen helper*
	Habitant	⟨Citizen⟩	600	*Aucklander, Brooklyner*
Secteur d'activité	Matière	⟨Discipline⟩	580	*art history*
	Branche	⟨Sector⟩	38 000	*life insurance, farming*
Organisation	Type d'organisation	⟨CompanyDescriptor⟩	23 800	*car manufacturer*
	Nom d'organisation	⟨Company⟩	516 000	*Fujitsu Siemens*
	Forme juridique d'entreprise	⟨LegalForm⟩	115	*ltd, inc, plc, AG, GmbH, s.a.r.l., LLC*
Lieu	Pays et Continent	⟨Nation⟩ ⟨Continent⟩	430	*South America, France, Germany, Europe*
	Ville	⟨City⟩	327 400	*'s-Gravenhage, Paris, Berlin*
Date	Mois	⟨Month⟩	24	*May, Jul., July*
	Jour de semaine	⟨DayOfWeek⟩	7	*Monday, Sunday*